Title

Identification of novel chondroitin proteoglycans in Caenorhabditis elegans: embryonic cell division depends on CPG-1 and CPG-2.

Journal

The Journal of cell biology, 173(6)

ISSN

0021-9525

Authors

Olson, Sara K
Bishop, Joseph R
Yates, John R
et al.

Publication Date

2006-06-01

DOI

10.1083/jcb.200603003

Peer reviewed

Identification of novel chondroitin proteoglycans in *Caenorhabditis elegans*: embryonic cell division depends on CPG-1 and CPG-2

Sara K. Olson,[1,2] Joseph R. Bishop,[1] John R. Yates,[4] Karen Oegema,[1,3] and Jeffrey D. Esko[1]

[1]Department of Cellular and Molecular Medicine, Glycobiology Research and Training Center, [2]Biomedical Sciences Graduate Program, School of Medicine, and [3]Ludwig Institute for Cancer Research, University of California, San Diego, La Jolla, CA 92093
[4]Department of Cell Biology, The Scripps Research Institute, La Jolla, CA 92037

Vertebrates produce multiple chondroitin sulfate proteoglycans that play important roles in development and tissue mechanics. In the nematode *Caenorhabditis elegans*, the chondroitin chains lack sulfate but nevertheless play essential roles in embryonic development and vulval morphogenesis. However, assignment of these functions to specific proteoglycans has been limited by the lack of identified core proteins. We used a combination of biochemical purification, Western blotting, and mass spectrometry to identify nine *C. elegans* chondroitin proteoglycan core proteins, none of which have homologues in vertebrates or other invertebrates such as *Drosophila melanogaster* or *Hydra vulgaris*. CPG-1/CEJ-1 and CPG-2 are expressed during embryonic development and bind chitin, suggesting a structural role in the egg. RNA interference (RNAi) depletion of individual CPGs had no effect on embryonic viability, but simultaneous depletion of CPG-1/CEJ-1 and CPG-2 resulted in multinucleated single-cell embryos. This embryonic lethality phenocopies RNAi depletion of the SQV-5 chondroitin synthase, suggesting that chondroitin chains on these two proteoglycans are required for cytokinesis.

Introduction

Vertebrates express at least 25 chondroitin sulfate proteoglycans (CSPGs), distinguished by the primary sequence of their core protein (Olson and Esko, 2004). Major groups include the aggrecan family of large, highly glycosylated proteoglycans present in connective tissues, the small leucine-rich proteoglycans, collagen α2IX, basement membrane proteoglycans (e.g., leprecan and perlecan), membrane bound proteoglycans (e.g., CD44, NG2, and phosphacan), and hybrid proteoglycans containing both chondroitin sulfate and heparan sulfate (syndecans). Secreted CSPGs create a hydrated matrix allowing for tissue expansion, and in cartilage they confer the ability to absorb compressive loading. In other tissues, they bind and help organize fibrillar collagens and can affect growth factor signaling pathways. Mutations affecting the core proteins or the enzymes involved in the assembly of the chondroitin chains result in developmental abnormalities in skin, cartilage, bone, tendon, eyes, brain, and the microvasculature (Iozzo, 1998; Hassell et al., 2002; Schwartz and Domowicz, 2002; Ozerdem and Stallcup, 2004; Schwartz and Domowicz, 2004; Goldberg et al., 2005; Marneros and Olsen, 2005; Niisato et al., 2005).

Each chondroitin proteoglycan (CPG) consists of a protein core and one or more covalently attached chondroitin chains. Assembly occurs in a stepwise manner, starting in the endoplasmic reticulum with translation of core proteins on membrane bound ribosomes and transfer of xylose to specific serine residues. Xylosylation exhibits specificity, occurring only at serine residues that have an adjacent glycine to the COOH-terminal side and one or more flanking acidic residues usually within eight amino acids (Esko and Selleck, 2002). After xylosylation, synthesis of a tetrasaccharide primer (-GlcAβ3Galβ3Galβ4Xylβ-O-Ser) takes place in the Golgi, followed by the polymerization of the chain by the alternating addition of *N*-acetylgalactosamine (GalNAc) and glucuronic acid (GlcA; $[GlcA\beta3GalNAc\beta4]_n$; Sugahara et al., 2003). In vertebrates, the chains undergo further modification by the addition of sulfate at C4 and C6 of GalNAc and C2 of GlcA residues (Habuchi, 2000). Additionally, an epimerase can convert a subset of GlcA residues to L-iduronic acid, which can be subsequently sulfated

Correspondence to Jeffrey D. Esko: jesko52@ucsd.edu

Abbreviations used in this paper: BEMAD, β-elimination followed by Michael addition with DTT; CPG, chondroitin proteoglycan; CSPG, chondroitin sulfate proteoglycan; DIC, differential interference contrast; dsRNA, double-stranded RNA; GalNAc, *N*-acetylgalactosamine; GlcA, glucuronic acid; MUDPIT, multidimensional protein identification technology; PH, pleckstrin homology.

The online version of this article contains supplemental material.

The Journal of Cell Biology, Vol. 173, No. 6, June 19, 2006 985–994
http://www.jcb.org/cgi/doi/10.1083/jcb.200603003

as well (dermatan sulfate). The pattern of sulfation varies in different tissues during development and in relation to age.

Interestingly, very little is known about the proteoglycans present in invertebrates, even in the well-studied nematode *Caenorhabditis elegans*. All of the components of the chondroitin biosynthetic machinery are completely conserved between *C. elegans*, mice, and humans, including formation of nucleotide sugar precursors and their import into the Golgi, assembly of the linkage region, and polymerization of the chain (Bulik et al., 2000; Berninsone et al., 2001; Hwang and Horvitz, 2002a,b; Yamada et al., 2002; Hwang et al., 2003a,b; Izumikawa et al., 2004). However, *C. elegans* lacks the sulfotransferases and the epimerase present in vertebrates, so the chains consist of unmodified GalNAc and GlcA residues (Yamada et al., 1999; Toyoda et al., 2000). In spite of its simplicity, genetic experiments demonstrate a crucial role for chondroitin in embryonic cell division and vulval morphogenesis (Herman et al., 1999; Bulik and Robbins, 2002; Hwang and Horvitz, 2002a,b; Hwang et al., 2003a,b; Mizuguchi et al., 2003; Izumikawa et al., 2004).

Although all of the enzymes required for chondroitin synthesis have been identified in *C. elegans*, no protein cores that bear the chondroitin chains have been described. To identify CPGs in *C. elegans*, we conducted searches using BLAST (Basic Local Alignment Search Tool) with all known mammalian CSPG core protein sequences but did not identify any obvious homologues. We therefore pursued a proteomics-based approach taking advantage of the chemical properties of the long, negatively charged chondroitin chains, a tagging method to modify the serine attachment sites, and mass spectrometry. We report the discovery of nine novel CPG proteins, none of which show homology to vertebrate CSPGs. Simultaneous depletion of two of these proteins, CPG-1/CEJ-1 and CPG-2, by RNAi results in defective cytokinesis during the first embryonic cell division. This phenotype is identical to that observed when the *sqv-5* chondroitin synthase was silenced by RNAi or by a loss-of-function mutation (Herman et al., 1999; Hwang and Horvitz, 2002b; Hwang et al., 2003b; Mizuguchi et al., 2003), suggesting that these two major chondroitin-carrying proteoglycans are essential for embryonic development.

Results

C. elegans expresses multiple CPGs

Proteoglycans were solubilized from a mixed-stage worm population by a combination of sonication and extraction with guanidine hydrochloride. After dialyzing the crude extract into urea, the proteoglycans were partially purified by stepwise elution from an anion-exchange column, using 0.2 M NaCl with urea to remove contaminating proteins and 1 M NaCl without urea to elute the proteoglycans (see Materials and methods). Treatment of the sample with sodium hydroxide released the chains from their core proteins by β-elimination reaction. The liberated chondroitin chains were generally <20 kD based on their elution by gel filtration chromatography relative to shark cartilage chondroitin sulfate (Fig. 1 A; Yamada et al., 1999), but they exhibited considerable heterogeneity. Treatment of the sample with chondroitinase ABC converted all of the material into nonsulfated disaccharides that eluted near the total volume of the column, confirming that the preparation contains predominantly chondroitin and very little heparan sulfate (Yamada et al., 1999; Toyoda et al., 2000).

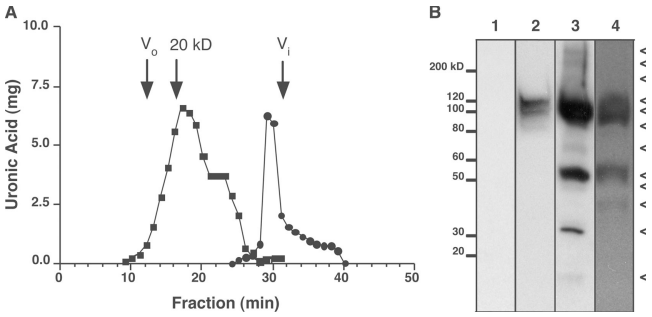

Figure 1. ***C. elegans* express multiple CPGs.** (A) Glycopeptides were treated with alkali to produce free glycans, which were then analyzed by gel filtration HPLC (see Materials and methods). Untreated sample eluted as two peaks (indicated by squares), whereas after chondroitinase ABC all the material eluted as disaccharides (indicated by circles). V_i, elution position of [1-^{3}H]glucose; V_o, elution position of blue dextran. Shark cartilage chondroitin sulfate (~20 kD) was run as a standard. (B) Crude extracts of adult worms (lanes 1–3) or embryos (lane 4) were digested with chondroitinase ABC (lanes 2–4) and separated by SDS-PAGE. Western blotting was performed with mAb 1-B-5, which recognizes a neoepitope generated by chondroitinase digestion. The undigested control (lane 1) demonstrates the requirement of the antibody for chondroitin stubs. Lanes 2 and 4, 10-μg protein; lanes 1 and 3, 40-μg protein. The carats indicate the position of protein bands that arose only after chondroitinase ABC digestion.

To estimate the number of CPGs, the crude *C. elegans* proteoglycans were digested with chondroitinase ABC, separated by SDS-PAGE, and analyzed by Western blotting with the 1B5 mAb that recognizes a neoepitope generated by chondroitinase digestion (Fig. 1 B). Several CPG core proteins ranging in mass from 10 to >200 kD were detected in whole worm (lanes 2 and 3) and embryo extracts (lane 4), whereas none were detected if chondroitinase ABC was omitted (lane 1). The major species had masses of ~118, ~110, and ~60 kD. Treatment of samples with PNGase F had no effect on the mobility of these bands, suggesting that none of the proteins contained Asn-linked glycans (unpublished data). *C. elegans* also synthesizes at least three heparan sulfate–bearing proteoglycans with approximate masses of 45, 28, and 22 kD based on digestion of extracts with heparinase followed by Western blotting with 3G10 mAb, which is specific for a heparan sulfate neoepitope generated by enzyme digestion (Hwang et al., 2003b). In vertebrates, some proteoglycans contain both chondroitin sulfate and heparan sulfate chains (e.g., syndecans 1 and 3). Simultaneous digestion of samples with chondroitinase ABC and heparin lyase II did not alter the intensity of any of the CPG bands or give rise to new ones reactive with mAb 1B5 or 3G10 (unpublished data). Thus, hybrid proteoglycans might not exist in *C. elegans*.

To identify the CPG core proteins detected by Western blotting, we searched the *C. elegans* genome for sequences related to any of the 25 known mammalian CSPG core proteins (Olson and Esko, 2004). No obvious homologues were found, with the exception of ORF Y47D3A.26, which resembled bamacan/SMC3 (Wu and Couchman, 1997). However, the coding sequence for this gene, as well as vertebrate bamacan, lacks

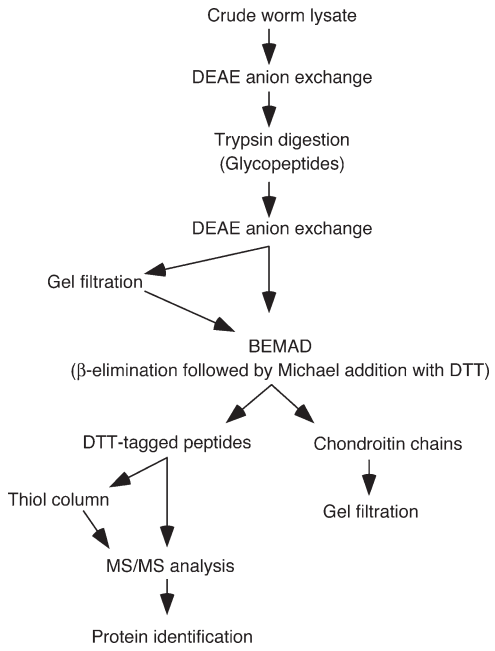

Figure 2. **Scheme for identifying CPGs in *C. elegans*.**

a signal peptide present in all known proteoglycans, consistent with the observation that most of bamacan is found in the nucleus in a nonglycosylated form (Ghiselli and Iozzo, 2000). Thus, a new method was needed to identify the various CPGs in *C. elegans*.

Fig. 2 describes the scheme used to identify *C. elegans* CPGs. After the DEAE chromatography step, samples were reduced, alkylated, and digested with trypsin. Peptides containing chondroitin chains (glycopeptides) were recovered by a second round of anion-exchange chromatography, which removed much of the contaminating protein from the sample. In some experiments, we used gel filtration to separate glycopeptides of large hydrodynamic size from contaminating peptides rich in acidic amino acids, which coeluted under the low salt conditions used to prepare the samples. Next, the samples were treated with sodium hydroxide, which caused β-elimination of the chains and the formation of dehydroalanine from serine residues. The dehydroalanine residues were then reacted with DTT, which adds across the double bond (Wells et al., 2002). In some experiments, the DTT-tagged peptides were further purified with a thiol column. Samples were then analyzed by mass spectrometry using multidimensional protein identification technology (MUDPIT) to identify peptides containing DTT (+167 D) or dehydroalanine (−18 D; Washburn et al., 2001; Wolters et al., 2001). The corresponding full-length proteins were identified by searching a *C. elegans* proteome database, parsing the data using the following criteria: (1) putative CPGs had to contain at least one glycosaminoglycan attachment site, defined as a modified serine residue flanked on the COOH-terminal side by glycine with one or more aspartate or glutamate residues near the site of chondroitin addition (Esko and Zhang, 1996), and (2) the proteins had to have a hydrophobic signal peptide, which would direct the protein into the secretory pathway, where chondroitin synthesis occurs. Nine tentative CPG core proteins were identified by this procedure and designated CPG-1 through -9 (Fig. 3).

Characteristics of the *C. elegans* CPGs

CPG-1 (C07G2.1a) and CPG-2 (B0280.5) have predicted masses of 62 and 54 kD, respectively (Fig. 3). CPG-1 was identified previously as CEJ-1 based on its reactivity with an antibody to a mammalian tight junction protein (Siddiqui, S.S., personal communication). It contains five putative chondroitin attachment sites based on amino acid sequence. One of these sites was confirmed by mass spectrometry based on presence of the DTT tag. In contrast, CPG-2 has 34 potential chondroitin attachment sites (four confirmed by mass spectrometry; Fig. 3). CPG-1 and -2 also have in common multiple peritrophin-A chitin binding motifs, defined by the arrangement of six cysteine

Protein Name (Accession Number)	Independent Verifications	Predicted Mass, kD (Apparent M_r)	Putative (Identified) Attachment Sites	Germline Enrichment
CPG-1/CEJ-1 (C07G2.1a)	2	62 (150)	5 (1)	Yes
CPG-2 (B0280.5)	6	54 (100-120)	34 (4)	Yes
CPG-3 (R06C7.4)	7	30 (60)	15 (0)	Yes
CPG-4 (C10F3.1)	4	84 (220)	35 (4)	ND
CPG-5 (C25A1.8)	3	27 (50)	1 (1)	Yes
CPG-6 (K10B2.3a)	2	28 (40)	1 (1)	Yes
CPG-7 (K09E4.6)	1	12 (ND)	11 (6)	ND
CPG-8 (K03B4.7a)	5	12 (ND)	6 (5)	No
CPG-9 (Y67D8C.8)	5	7 (ND)	4 (2)	ND

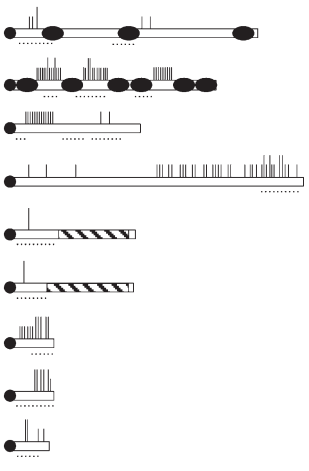

Figure 3. ***C. elegans* CPGs identified by BEMAD/MUDPIT.** CPGs were purified from worm extracts and identified by mass spectrometry analysis. Nine independent purifications resulted in identification of nine CPGs, most of which were identified in multiple runs. The predicted masses are based on amino acid sequence, and the apparent masses are based on SDS-PAGE migration of recombinant proteins expressed in COS-7 cells after digestion with chondroitinase ABC and reduction. Putative glycosylation sites (short vertical lines) consist of Ser-Gly dipeptides flanked by one or more acidic amino acids. Identified sites (long vertical lines) represent serine residues modified with DTT by the BEMAD method (see Materials and methods). Genes enriched for germline expression were identified by Reinke et al. (2000). Schematic drawings of each CPG are shown. Gray circles indicate signal peptides, black ovals are peritrophin-A chitin binding domains, diagonally hatched boxes identify C-type lectin domains, and dotted lines below the protein indicate peptide coverage discovered by mass spectrometry.

residues that can form three disulfide bridges in a characteristic pattern (Wright et al., 1991; Beintema, 1994; Venegas et al., 1996; Merzendorfer and Zimoch, 2003). In CPG-1, two peritrophin-A domains lie in the NH_2-terminal half of the protein and one toward the COOH-terminal end, whereas in CPG-2 the six peritrophin-A domains lie between sites predicted to be glycosylated (Fig. 3). Additionally, the COOH-terminal half of CPG-1 contains many Thr, Val, and Pro residues, resembling a segment of Muc-2, a membrane mucin present in vertebrates.

CPG-3 (R06C7.4) was not originally identified by the DTT-tagging method or by the presence of dehydroalanine (Wells et al., 2002) but rather by 5′-RACE using PCR primers designed against the repetitive SG region of *cpg-2*. Reevaluation of the mass spectral data revealed that peptides from CPG-3 were present in nearly all datasets (Fig. 3). Based on primary sequence, CPG-3 contains 15 putative chondroitin attachment sites and shows 21% sequence identity to CPG-2, mainly in the putative glycosylated region. The lack of mass spectral data for peptides that contain putative attachment sites in CPG-3 may have been caused by the failure of trypsin to gain access to and cleave the protein in heavily glycosylated regions.

CPG-4 (C10F3.1) has the largest predicted mass of the *C. elegans* CPGs at 84 kD. The majority of its 35 predicted glycosaminoglycan attachment sites reside in the COOH-terminal half of the protein, of which four sites were confirmed by DTT modification (Fig. 3).

CPG-5 (C25A1.8, CLEC-87) and CPG-6 (K10B2.3a, CLEC-88) are highly homologous, showing 67% identity and 90% similarity. CPG-5 and -6 contain only one putative chondroitin site toward the NH_2 terminus, which was identified in both proteins by DTT modification. Both proteins contain a C-type lectin domain in the COOH-terminal half of the protein, like the aggrecan family of vertebrate CSPGs (Zelensky and Gready, 2003). This motif therefore classifies *cpg-5* and *-6* as members of the *clec* (C-type lectin) gene class. Proteins containing C-type lectins are common in *C. elegans* (88 members have been found thus far), but their glycan binding specificity has not been determined.

CPG-7 (K09E4.6) and CPG-8 (K03B4.7a) have predicted masses of only ∼12 kD (Fig. 3). CPG-7 contains 11 putative chondroitin attachment sites, whereas CPG-8 contains 6 sites. In both cases, five of the sites were confirmed by mass spectrometry. Of the multiple peptides identified in CPG-8 samples, some contained a single modified site at either residue 61 or 63 (the tandem mass spectrometry data did not discriminate which site actually carried the chain). Interestingly, other peptides were modified at both serine residues. A similar phenomenon was observed at residues 84 and 88. CPG-9 (Y67D8C.8) has a predicted mass of only 7 kD, and mass spectrometry confirmed two of four attachment sites.

Interestingly, none of the CPGs identified using the scheme in Fig. 2 showed homology to vertebrate CSPGs. Furthermore, no homologues were found in searches of the *D. melanogaster* genome or *Hydra vulgaris* EST and genomic databases, suggesting that the core proteins have evolved independently in different organisms. In contrast, homologues for each *C. elegans* CPG were present in *Caenorhabditis briggsae*, another free-living nematode.

The glycosylation sites of many *C. elegans* CPGs appeared to follow an "SGEX" motif, where X was usually G, A, T, or S (Fig. 4). BLAST searches using this sequence motif identified CPG-2, -3, -4, and -7 but no other proteins in *C. elegans* or vertebrates. *D. melanogaster* contains one protein containing this motif (CG6048-PA), but it has not yet been shown to contain a chondroitin chain. The aggrecan family of CSPGs in vertebrates also contains repetitive SG sequences within domains densely substituted with chondroitin chains, but the intervening sequences differ from those found in *C. elegans* CPGs.

C. elegans core proteins carry chondroitin sulfate chains after expression in mammalian cells

To confirm that the *C. elegans* CPGs identified by mass spectrometry could serve as a scaffold for chondroitin synthesis, the cDNA sequences were subcloned into a mammalian expression vector with a COOH-terminal Myc tag and expressed in COS-7 cells.

```
CPG-1  TTTLDGSGDGSGADNGFVSGADAVAI    57
        ECTNGSGNDEGSADETTPESSGEMPY  290

CPG-2  EGSGESSGETSGEGSGESSGEASGEGSGEASGEGSGEASGEGSGEASGEGSGSGEE  132
           DGSGETSGEGSGEASGEASGEGSGEASGESSGQGSGEASGEGSGELE        244
              ETSGEESGEASGEQSGEGSGEASGEASGESSGEGSGVEE             399

CPG-3  EFSGDSSGEISGESSGEASGEASGEASGEASGEASGESSGETSGESSGDEETSGEGSGEEGSGDT 99
       GDSGCGSGKCIG  207

CPG-4         ESSGSSGEMSGDGSDNEASGEGSGEYDASGSSGDNSGEFNSSGSSGEA  567
            ESSGEFSGSSGEGSGDTASSDTSIDDKSIIRSGEGSAESVSEILQEASGED  641
       DHSGFGESSGSSGESIELRDSGEG                                  681
       ESSGDYEFSGSSNESIEQSKEGSAASIYEILQAASGED                    768
             DASGSSGDNSGDFNSSGSSGEASGVGESSGSEDQGSGNY             724
                                       ESSGSEDQGSGNY             584

CPG-5   ELPEASGEAPET  37

CPG-6   DIEDASGETPGI  33

CPG-7  ESSGEGSGESSGEKPVVESSGEGSGESSGDKEAVEASGEGSGEGSGDATLESSGEGSGES  91

CPG-8    DGSGSGESPAEEQLRRVVRDVDEEASGEGSGAAEVTSVPVRFVRSVDAEGSGSGSDE  115

CPG-9  DLEGSGSGDVSTDAKEAILNAQTLLDAVSSDGSGADVEASGED  63
```

Figure 4. **Chondroitin attachment sites in CPGs.** Primary sequences of CPG peptide segments containing chondroitin attachment sites are shown using the single letter code for each amino acid. Underlined SG dipeptides refer to attachment sites in which the serine residues were modified by DTT or converted to dehydroalanine residues during BEMAD.

Conditioned media was collected 48 h after transfection, purified by anion-exchange chromatography, treated with chondroitinase ABC, and Western blotted with an antibody to Myc (Fig. 5). CPG1-6 yielded reactive protein bands that did not appear in the absence of chondroitinase ABC digestion. Attempts to express CPG7-9 in COS cells have thus far been unsuccessful, possibly because of the small size of the protein core. None of the proteoglycans migrated at the molecular mass predicted by their primary sequence. Furthermore, a comparison of the bands in Fig. 1 B to the calculated masses in Fig. 3 did not demonstrate obvious correspondence. However, the migration pattern of CPG-2 expressed in COS cells appears similar to the 120-kD core in the crude proteoglycan preparation (Fig. 1 B). Likewise, CPG-3 migrates at 55–60 kD, the same position as a prevalent band in the worm extract. These observations suggest the presence of higher order structures in the core proteins or that the chondroitin stub oligosaccharides remaining after chondroitinase digestion altered their migration. Mammalian CSPGs also tend to migrate aberrantly, often at twice their predicted size (Bellin et al., 2002).

Simultaneous RNAi of *cpg-1* and *-2* leads to penetrant embryonic lethality

Expression of *cpg-1*, *-2*, *-3*, *-5*, and *-6* is enriched in the *C. elegans* germline, suggesting possible activity in germline development and progression, spermatogenesis, oogenesis, or embryogenesis (Reinke et al., 2000; Fig. 3). However, genome-wide RNAi screens have not reported phenotypes for any of the *cpg* genes, suggesting either functional redundancy or subtle phenotypes (www.wormbase.org). Lee and Schedl (2001) reported that GLD-1 acts as a translational repressor of several mRNA targets in the germline, including *cej-1* and B0280.5, which were identified in our proteomic screen as *cpg-1* and *-2*, respectively. They also reported that RNAi of each gene alone had no phenotype, whereas simultaneous depletion of both genes resulted in embryonic lethality. Because this phenotype is reminiscent of the embryonic phenotype resulting from inhibiting the SQV-5 chondroitin synthase, we decided to analyze CPG-1 and -2 in greater detail.

To confirm that CPG-1 and -2 contain chondroitin chains in vivo, worms were fed bacteria expressing double-stranded RNA (dsRNA) directed against *cpg-1* or *-2*, and samples were analyzed by Western blotting. No differences were initially observed in CPG-1–depleted extracts (Fig. 6 A), but CPG-2–depleted extracts showed a dramatic reduction of the major band at ~120 kD, as well as the bands at ~80 and ~45 kD (Fig. 6 A, arrowheads), suggesting that all three were related to each other (the predicted mass of the protein was 54 kD; Fig. 3). The simultaneous reduction of these three bands was confirmed by using a second construct targeting a different region of *cpg-2* (unpublished data). The RNAi effect was specific, as the other major core protein band at ~60 kD remained unaltered.

CPG-1 and -2 have in common multiple peritrophin-A chitin binding domains (Fig. 3). To test the functionality of these domains, crude worm extract was incubated with chitin beads. Four CPG bands were able to bind chitin, including those at 120, ~80, and 45 kD (Fig. 6 B, filled arrowheads), as well as a new band at ~150 kD (Fig. 6 B, open arrowhead). Chitin binding appeared to be selective, as the major band at 60 kD did not interact with the resin. *cpg-2* RNAi treatment diminished the three bands at 120, ~80, and 45 kD, confirming the presence and functionality of one or more peritrophin-A chitin binding domains in these proteoglycans (Fig. 6 C, filled arrowheads). The band migrating at ~150 kD was present in vector and CPG-2–depleted extracts but absent in extracts depleted of CPG-1 and CPG-1/CPG-2 (Fig. 6 C, open arrowhead). Recombinant CPG-1 expressed in COS-7 cells had a similar mass (Fig. 5), suggesting the 150-kD band represents CPG-1. Thus, both CPG-1 and -2 behave as CPGs in vivo and have functional chitin binding domains.

To study the function of CPG-1 and -2, we compared the effect of silencing their expression by RNAi to depletion of SQV-5, which encodes the chondroitin synthase. Brood sizes of *cpg-1(RNAi)*, *cpg-2(RNAi)*, and c*pg-1/cpg-2(RNAi)* animals were comparable to uninjected worms or worms injected with

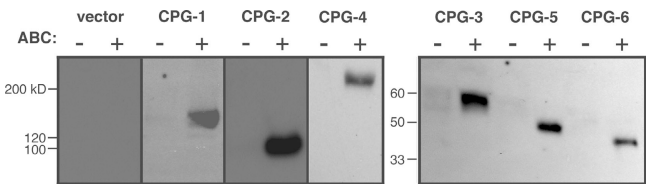

Figure 5. **CPG-1 through -6 behave as CSPGs in COS-7 cells.** CPG-1 through -6 were expressed as Myc-tagged recombinant proteins in COS-7 cells. Proteoglycans from conditioned media were purified by anion-exchange chromatography (see Materials and Methods). Samples were digested with chondroitinase ABC (+) or left untreated (−), separated by SDS-PAGE, and Western blotted with an anti-Myc mAb.

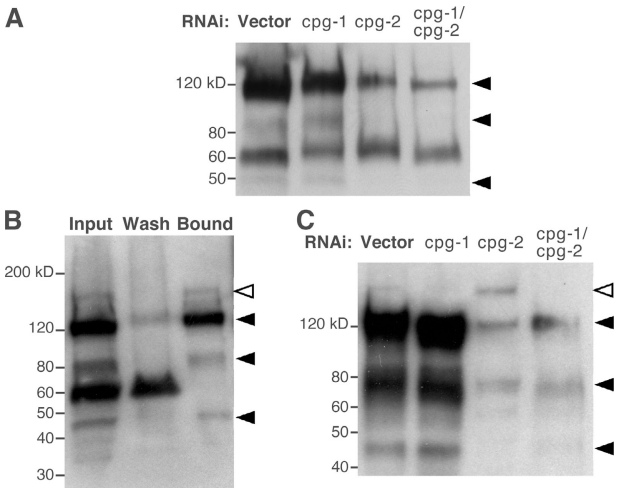

Figure 6. **CPG-1 and -2 bind chitin and exist as CPGs in vivo.** (A) Proteoglycan extracts of vector and RNAi-treated animals were digested with chondroitinase ABC, analyzed by SDS-PAGE, and Western blotted with the 1B5 mAb that recognizes the chondroitin stub remaining after enzyme digestion. *cpg-2(RNAi)* reduced the bands at 120, ~80, and ~45 kD (arrowheads). (B) 20 µg of total worm extract (input) was incubated with chitin beads. Four bands specifically bound chitin (arrowheads), but the 60-kD band did not. (C) Proteoglycan extracts from RNAi-treated animals were enriched by affinity chromatography on chitin beads. *cpg-1(RNAi)* reduced the band at ~150 kD (open arrowhead), whereas *cpg-2(RNAi)* reduced the bands at 120, ~80, and 45 kD (filled arrowheads).

buffer alone (Table I; $P > 0.05$ by one-way analysis of variance). Depletion of CPG-1 or -2 alone also had no effect on viability, as 96–99% of the embryos hatched into healthy larvae that grew into normal fertile adults (Table I). However, simultaneous depletion of CPG-1 and -2 had a synergistic effect, resulting in penetrant embryonic lethality. Whether this synergy is due to the functional redundancy of CPG-1 and -2, the partially penetrant nature of RNAi, other functions of the proteoglycans besides carrying chondroitin chain, or the cross-reactivity of *cpg-2* dsRNA with other targets remains to be determined. *sqv-5(RNAi)* treatment also had no effect on egg production or laying but had a hatching phenotype similar to c*pg-1/cpg-2(RNAi)* (95% of progeny failed to hatch). To look for genetic interactions between *sqv-5* and *cpg-1* or *-2,* a heterozygous *sqv-5(n3611)* mutation was compounded with single c*pg-1(RNAi)* or *cpg-2(RNAi)*. However, no combination affected brood sizes or embryonic viability (unpublished data), possibly because the heterozygous mutation did not reduce chondroitin levels sufficiently to see the effect of depleting a single CPG.

Table I. **Brood size and viability in RNAi-treated worms**

RNAi treatment	Brood size[a]	Range	Viability
			%
Uninjected	133 ± 43	75–197	98.9
Buffer	120 ± 45	65–178	99.6
cpg-1	95 ± 34	64–184	95.9
cpg-2	144 ± 42	87–202	99.2
cpg-1/cpg-2	129 ± 56	62–203	0.0
sqv-5	77 ± 22	22–114	4.8

The number of embryos and hatched L1 larvae were assessed 24 h after placing control worms or worms injected with dsRNA on individual plates. Viability was determined >24 h after the initial brood size was counted and was calculated as the number of hatched larvae divided by the total number of embryos laid. $n = 6–11$.
[a]One-way analysis of variance showed no significant difference between brood sizes. $P > 0.05$.

RNAi of *cpg-1* and *-2* results in a cytokinesis defect

To determine the underlying cause for embryonic lethality, we looked more closely at the first cell division, where defects have been reported in *sqv-5* chondroitin synthase mutants (Hwang et al., 2003b). To better visualize the plasma membrane, a strain expressing GFP fused to a pleckstrin homology (PH) domain derived from mammalian phospholipase-C was injected with dsRNA against *cpg-1*, *cpg-2*, *cpg-1/cpg-2*, or *sqv-5*. The PLC1δ1 PH domain binds to phosphatidylinositol 4,5 bisphosphate located specifically on the plasma membrane and is an excellent probe for monitoring the changes in cell shape that accompany cytokinesis (Audhya et al., 2005). The depleted embryos were simultaneously imaged by differential interference contrast (DIC; Fig. 7, top) and spinning-disc confocal microscopy (Fig. 7, bottom). Embryos were imaged in utero because depletion of either SQV-5 or both CPG-1 and -2 resulted in embryos that were fragile and osmotically sensitive. Depletion of CPG-1 or -2 alone had no effect on fertilization, membrane ruffling, pseudocleavage, pronuclear meeting and rotation, karyokinesis, and cytokinesis (compare Video 2 to buffer-injected worms in Video 1, available at http://www.jcb.org/cgi/content/full/jcb.200603003/DC1). Initiation of the cleavage furrow occurred normally (Fig. 7 D, bottom), and two separate daughter cells routinely appeared (Fig. 7 E). A four-celled embryo was distinctly seen at the end of filming, demonstrating the fidelity of the second round of cell division (Fig. 7 G). In contrast, embryonic development was severely perturbed in embryos simultaneously depleted of CPG-1 and -2 (Video 3). In wild-type embryos, fertilization triggers the oocyte pronucleus to undergo two rounds of meiotic segregation, which produces two polar bodies that are extruded from the embryo by small cytokinesis-like events. In CPG-1/CPG-2–depleted embryos, fertilization occurred normally but polar body extrusion failed, resulting in extra nuclear material that remained in the embryo (Fig. 7, I and J). The pronuclei migrated and the complex of the pronuclei and centrosomes rotated onto the long axis of the cell as in wild type, but the membrane ruffling that normally precedes and accompanies these events was absent, as was the space between the embryonic plasma membrane and the eggshell that is normally present during cell division (compare Fig. 7, C [bracket] and J [solid white lines]). Spindle formation appeared normal and the chromosomes appeared to segregate, but the cleavage furrow failed to form and ingress to introduce a new cell surface between the segregated chromosomes (Fig. 7, K and L). Without the barrier of a new cell membrane, the daughter nuclei fused (Fig. 7 M) and repeatedly attempted cell division without cytokinesis (Fig.7 N). An essentially identical phenotype was observed after depletion of SQV-5 (Fig. 7, O–U; and Video 4). Since depleting CPG-1 and -2 mimics the defect in embryonic cytokinesis by depletion of the SQV-5 chondroitin synthase, we conclude that CPG-1 and -2 are two functionally important CPGs required at this early developmental stage.

Discussion

A proteomic approach for identifying novel proteoglycans

We report here the identification of a group of novel CPGs in *C. elegans*. The identification scheme used a combination of conventional methods to purify proteoglycans and glycopeptides based on the polyanionic character and size of the chondroitin chains coupled with mass spectrometry. Additional selection criteria ensured that the majority of identified proteins had the properties of a proteoglycan. These included chondroitin attachment sites modified with DTT, the presence of a serine residue followed by glycine and one or more nearby acidic residues, and a signal peptide. Although DTT addition was not stoichiometric, the β-elimination step introduced an alternate method to identify the protein via the characteristic mass difference between dehydroalanine and serine (Wells et al., 2002). Attempts to optimize the DTT-addition reaction have not yet yielded higher efficiency, and other methods are needed to improve this step to better assess the stoichiometry of chondroitin attachment. Similarly, adjustment of the proteolysis conditions might improve peptide coverage, as the absence of trypsin

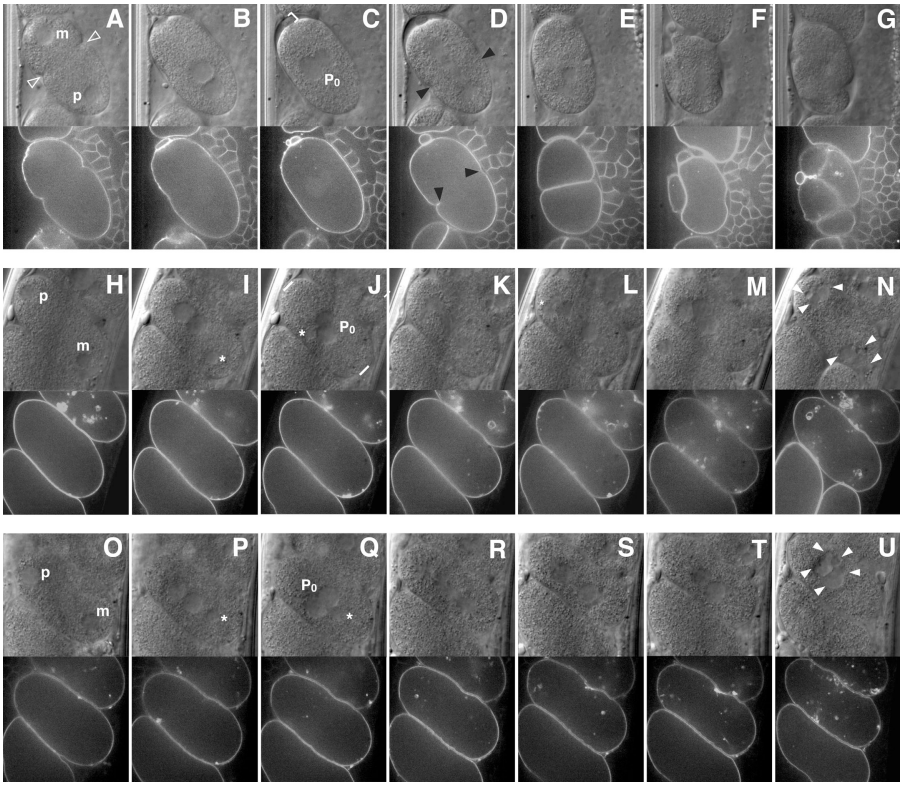

Figure 7. **CPG-1 and -2 are required for embryonic cell division.** The first two cell divisions of embryos depleted of CPG-2 (A–G, see Video 2, anterior toward the top left, available at http://www.jcb.org/cgi/content/full/jcb.200603003/DC1), CPG-1/CPG-2 (H–N, see Video 3, anterior toward the bottom right), and SQV-5 (O–U, see Video 4, anterior toward the bottom right) were analyzed by DIC Nomarski microscopy (top) and membrane-GFP fluorescence (bottom). (A, H, and O) Pseudocleavage initiates in the *cpg-1* single RNAi (open arrowheads), but not in the *cpg-1/cpg-2* double or *sqv-5* RNAi embryos. The maternal (m) and paternal (p) pronuclei are indicated. (B, I, and P) Pronuclear fusion proceeds normally in single, double, and *sqv-5* RNAi embryos. However, extrusion of the second polar body fails in doubly depleted and *sqv-5* embryos (indicated by asterisks). (C, J, and Q) Rotation of the fused nuclei. (D, K, and R) Initiation of the first cell division. The cleavage furrow is seen in *cpg-1(RNAi)* embryos (black arrowheads). (E, L, and S) Cytokinesis is complete in *cpg-1(RNAi)* embryos, forming two daughter cells. Nuclear division appears normal in *cpg-1/cpg-2(RNAi)* and *sqv-5(RNAi)* embryos, but cytokinesis did not initiate, causing the daughter nuclei and polar body to rejoin (M and T). Cell division is attempted again (N and U), resulting in multiple nuclei in a single-celled embryo (white arrowheads). The second cell division (F and G) proceeds normally in *cpg-1(RNAi)* embryos.

cleavage sites and the potential inhibitory effects of closely spaced chondroitin chains could interfere with the analysis. Regardless of its limitations, the current protocol led to the identification of nine proteins that had the hallmarks of CPGs, and six served as substrates for chondroitin sulfate assembly when expressed in mammalian cells. The simplicity of the identification strategy suggests that it should be generally applicable to other organisms, as well as tissues, cells, and secretions from vertebrates for which little information about the proteoglycan composition exists.

Interestingly, none of the *C. elegans* CPGs identified in this study had predicted membrane spanning domains or consensus sites for glycosphosphatidylinositol-anchor attachment. *C. elegans* expresses homologues of the membrane heparan sulfate proteoglycan syndecan, which has a single membrane spanning segment, and the glycosphosphatidylinositol-anchored heparan sulfate proteoglycan, glypican. Because vertebrates express multiple membrane CSPGs (e.g., NG2, CD44, and phosphacan; Olson and Esko, 2004), our failure to identify membrane CPGs in *C. elegans* could indicate either their absence or that they represent only minor components of the proteoglycan population. Heparan sulfate proteoglycans also were not detected, most likely because *C. elegans* expresses ~250 times less heparan sulfate than chondroitin (Yamada et al., 1999; Toyoda et al., 2000). Thus, it may be necessary to modify the procedure to selectively enrich membrane proteins and increase sample size to detect rare proteoglycans.

The lack of homology between *C. elegans* CPGs and vertebrate CSPGs and the observation that none of the CPGs have homologues in two other invertebrates (*D. melanogaster* and *H. vulgaris*) raises interesting questions about the evolution of these molecules. All of the core proteins identified in this study have sequence motifs required for initiation of glycosaminoglycan biosynthesis similar to those found in vertebrates (Esko and Zhang, 1996). Furthermore, *C. elegans* expresses orthologues of all of the vertebrate enzymes required for assembly of the linkage region tetrasaccharide (xylosyltransferase, galactosyltransferases I and II, and glucuronyltransferase I; Bulik et al., 2000; Hwang et al., 2003a), as well as the chondroitin polymerizing system (Hwang et al., 2003b; Mizuguchi et al., 2003), including the newly discovered chondroitin polymerizing factor (Izumikawa et al., 2004). These findings indicate that the mechanism of chain initiation and polymerization evolved early in metazoans and has been maintained by strong selection. In contrast, the core proteins on which the chains assemble have continued to evolve, presumably to serve specialized functions. For example, CPG-1 and -2 have functional chitin binding domains that could interact with the nematode eggshell, whereas no vertebrate proteoglycans contain this motif and vertebrates lack chitin as a structural component. The presence of chitin in insect exoskeletons and peritrophic matrices that line the gut suggests that orthologues of these CPGs might exist in *D. melanogaster* (Merzendorfer and Zimoch, 2003), but BLAST searches have not yet uncovered them. Additional studies using techniques like the one reported here are needed to identify the proteoglycans present in *D. melanogaster* and other organisms, which in turn might help define the evolution of proteoglycans during the metazoan expansion.

CPG-1/CEJ-1 and CPG-2 play essential roles in embryonic cell division in *C. elegans*

Our preliminary analysis of RNAi-depletion experiments indicated that silencing *cpg-1* or *-2* had no effect on embryogenesis or morphogenesis. However, depletion of both genes resulted in a strong embryonic phenotype characterized by failure of polar body extrusion after fertilization, loss of membrane ruffling preceding pronuclear fusion, and failure to initiate the cleavage furrow before cytokinesis. The necessity of depleting both CPG-1 and -2 to uncover a phenotype suggests functional redundancy of the proteins, though the only commonality is the presence of chitin binding domains and at least one chondroitin chain (Fig. 3). An alternative explanation is that depletion of each protein by RNAi may not have been fully penetrant, and the phenotype after double depletion could be the result of more extensive depletion of chondroitin. However, this explanation seems unlikely because CPG-1 is a minor component and had to be enriched by chitin affinity chromatography to detect its presence. Instead, the requirement for simultaneous depletion of both proteoglycans could be due to other functional aspects of the CPGs (e.g., their ability to bind chitin or the presence of other glycans). Generation of deletion mutants will help resolve this issue.

The strong embryonic lethal phenotype seen with *cpg-1/cpg-2(RNAi)* resembles that of the *sqv-5* chondroitin synthase mutant described by Hwang et al. (2003b). How the CPGs mediate these effects is unknown. One of the roles of CPGs may be to fill the space between the eggshell and the embryo, and the presence of a high concentration of polyanions and their counterions might cause sufficient hydrostatic pressure to aid in formation of the extraembryonic space and subsequent ingress of the cleavage furrow. Thus, the loss of CPGs in the perivitelline space surrounding the embryo could result in collapse of the extracellular space and apposition of the plasma membrane to the eggshell (Fig. 7). Interestingly, chondroitin has been shown to be present on both the embryonic cell surface and the eggshell, suggesting that it may function at both locations (Sugahara et al., 2003). This leads to an alternative explanation: CPG-1 and -2 could act as structural elements of the eggshell or might bridge chitin polymers in the eggshell with other components of the embryonic plasma membrane that result in transmembrane signaling to cytoskeletal components involved in cytokinesis. Regardless of which hypothesis is correct, we have shown that CPG-1 and -2 play a crucial role in cytokinesis, as their simultaneous depletion results in multinucleated single cell embryos.

The observation that the regulatory protein GLD-1 may translationally regulate *cpg-1* and *-2* transcripts indicates a sophisticated level of control over proteoglycan expression during early embryogenesis. GLD-1 is localized to the germline cytoplasm (Jones et al., 1996) and translationally regulates the expression of a large number of genes by interacting with mRNA 3′UTR sequences (Lee and Schedl, 2001). It also protects some transcripts from nonsense-mediated mRNA decay by binding to 5′UTR sequences (Lee and Schedl, 2004). One example is *gna-2*, which encodes the enzyme glucosamine 6-phosphate *N*-acetyltransferase. This enzyme plays an essential role in the biosynthesis of the nucleotide sugar UDP-*N*-acetylglucosamine (UDP-GlcNAc). The biosynthesis of a variety of glycans depends on UDP-GlcNAc, including chitin (a polymer of GlcNAcβ1,4) and chondroitin by way of UDP-GalNAc. Interestingly, a null allele of *gna-2(qa705)* exhibits maternal effect lethality and multinucleated embryos (Johnston, W., and Dennis, J., personal communication). The similar phenotype of *cpg-1/cpg-2(RNAi)*, *sqv-5(RNAi)*, and *gna-2* mutants may indicate a common mechanism of action through formation of CPGs. Posttranscriptional regulation of CPG assembly in *C. elegans* via GLD-1 is reminiscent of the regulation of enzymes involved in heparan sulfate formation in vertebrates (Grobe and Esko, 2002) and may indicate the general importance of translational control over proteoglycan assembly during development.

Materials and methods

C. elegans maintenance

The OD58 strain (provided by A. Audhya, University of California, San Diego, La Jolla, CA) carries a PH membrane domain of phospholipase-C fused in-frame to a GFP reporter (*unc119; pAAI pie-1::GFP-PH (PLC1δ1 PH); unc-119*) under the control of the *pie-1* promoter (Audhya et al., 2005). All other strains were obtained from the Caenorhabditis Genetics Center and cultivated as described previously (Brenner, 1974). Worm vectors were provided by A. Fire (Stanford University School of Medicine, Stanford, CA).

In silico analysis

BLAST searches were performed in the National Center for Biotechnology Information database (http://www.ncbi.nlm.nih.gov) or WormBase (http://wormbase.org) using annotated human or mouse CSPG core protein sequences. The presence of signal peptides was confirmed with PSORT II (http://psort.hgc.jp/) and SignalP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP).

Biochemical purification

Worm extracts were prepared from 20-g batches of a mixed-stage N2 (Bristol) population by sonication in 50 mM sodium acetate buffer, pH 6. Material was extracted for 48 h at 4°C in three volumes of solution containing 4 M guanidine-HCl, 0.1 M NaCl, 0.3% CHAPS, 50 mM sodium acetate, pH 6, and protease inhibitors (1 μg/ml leupeptin, 1 mM PMSF, and 1 μg/ml pepstatin A; Esko, 1993). The extract was dialyzed three times against 40 volumes of 6 M urea, 0.1 M NaCl, and 50 mM sodium acetate, pH 6. Insoluble material was removed by low-speed centrifugation and filtration through filter paper (No. 1; Whatman). The concentration of protein was assayed by the Bradford method (Bio-Rad Laboratories).

Proteoglycans were purified from worm extract (250 mg protein) by anion-exchange chromatography (0.2 M – 1 M NaCl, DEAE-Sephacel) and desalted by gel filtration (Bame and Esko, 1989). Partially purified material was treated with trypsin (Wolters et al., 2001), and glycopeptides were further purified by an additional pass over DEAE-Sephacel and in some experiments by gel filtration HPLC (TSK-2000 column [Tosoh Bioscience]; 1 M NaCl in 10 mM KH_2PO_4, pH 6). Fractions containing uronic acid (Esko and Manzi, 1996) were pooled and desalted by gel filtration. Glycosaminoglycan chains were removed by β-elimination followed by Michael addition with DTT (BEMAD; Wells et al., 2002). In brief, glycopeptides were incubated for 3 h at 50°C in 20% ethanol, 1% triethylamine, 10 mM DTT, and 0.1% NaOH. The reaction was quenched by adjusting the sample to 1% trifluoroacetic acid, and the peptides were purified by reverse-phase chromatography on a C18 Sep-Pak cartridge (Waters Corporation) by elution with 70% acetonitrile in 0.1% trifluoroacetic acid. Samples were dried before further analysis. In some experiments, the samples were passed over a thiol-Sepharose column to enrich for DTT-modified peptides (Wells et al., 2002).

Mass spectrometry

Tagged peptides were identified by MUDPIT (Washburn et al., 2001; Wolters et al., 2001; MacCoss et al., 2002; Washburn et al., 2002). Peptides were eluted stepwise from a biphasic capillary column made of strong cation-exchange resin coupled to a reverse-phase resin directly into

a tandem mass spectrometer. The identities of the DTT-tagged peptides were determined by searching the tandem mass spectra against a *C. elegans* proteome database using SEQUEST software and a computer array. The unique mass signature imparted by DTT (+167 D) and unmodified dehydroalanine residues (−18 D) were used to determine the sites of chondroitin addition.

Recombinant protein expression

C. elegans cDNA was prepared from total RNA with the SuperScript III First-Strand kit (Invitrogen). *cpg-1* through *-9* were amplified from *C. elegans* cDNA with primers that included the start codon and the penultimate codon and restriction sites to subclone the products in-frame into the pcDNA3.1(-)MycHis B vector (Invitrogen). Expression constructs were transfected into COS-7 cells with Lipofectamine (Invitrogen) following the manufacturer's instructions. Media was harvested 48 h later and purified over DEAE-Sephacel as described previously (Bame and Esko, 1989). COS-7 cells were cultured in Dulbecco's modified Eagle's medium (CellGro) supplemented with 10% fetal bovine serum, 100 U/ml penicillin G, and 100 μg/ml streptomycin sulfate.

RNAi

C. elegans cDNA was amplified by PCR with primers engineered to contain T7 (forward primer) or T3 (reverse primer) bacterial promoter sequence and base pairs 1–649 for *cpg-1* and 1–*775*, *754*–1572, and 1190–1572 for *cpg-2*. dsRNA was generated with the Megascript T7 and T3 transcription kits (Ambion) according to the manufacturer's instructions. OD58 L4 or young adult worms were injected with dsRNA and allowed to recover at 16°C or 20°C. To count brood sizes, worms were transferred 24 h after injection to individual plates, and 1 d later the number of eggs and hatched L1 larvae were counted (Maddox et al., 2005). Viability was measured 24–36 h later by counting the number of hatched larvae and unhatched embryos. Percentage of viability was calculated as number of hatched progeny divided by the total number of eggs laid.

Early embryonic cell division was assessed 24 h after injection. Embryos were filmed in utero because *cpg-1/cpg-2(RNAi)* and *sqv-5(RNAi)* embryos were fragile and osmotically sensitive. Injected animals were anesthetized with 1 mM levamisole in M9 buffer, mounted on an agarose pad, and filmed as described previously (Maddox et al., 2005). Images were acquired on a DeltaVision deconvolution microscope (Applied Precision) equipped with a charge-coupled device camera (CoolSnap; Roper Scientific) at 20°C. Images were acquired using a 100×, 1.3 NA U-Planapo objective (Olympus) with a 2 × 2 binning and a 480- × 480-pixel area. DIC and GFP images were acquired at 10-s intervals by sequentially rotating the analyzer and GFP filter set into the light path. Illumination was attenuated with a 10% neutral density filter. Images were analyzed with MetaMorph software (Universal Imaging Corp.).

To determine the effect of RNAi on proteoglycan expression, cDNA sequences for *cpg-1* and *-2* described earlier in this section were subcloned into vector pL4440, which carries dual T7 promoter sites (Fire Lab Vector kit) and drives the formation of dsRNA. The expression vector was transformed into the HT115 bacterial strain, which was then fed to L1 larval worms as described previously (Kamath et al., 2001). After 48–60 h, protein extracts were prepared by sonication of either whole worms or embryos and subjected to SDS-PAGE and Western blotting as described in the next section.

Chitin binding experiments were performed with chitin beads (New England Biolabs, Inc.) according to the manufacturer's instructions. Unbound protein (including the wash buffer) was desalted by gel filtration before analysis. Bound material was analyzed by boiling chitin beads in SDS buffer and loading the sample directly onto the gel.

Western blotting

Proteoglycan samples (10 μL) were digested with 10 mU chondroitinase ABC (Seikagaku) and/or 2.5 mU heparin lyase II (Sigma-Aldrich) for 3–5 h at 37°C. After reduction with β-mercaptoethanol and alkylation with iodoacetamide, samples were analyzed by SDS-PAGE. CPG core proteins were detected with 1B5 chondroitin stub mAb (1:1,000; Seikagaku) followed by goat anti–mouse secondary antibody (1:2,000; Bio-Rad Laboratories). Recombinant proteins expressed in animal cells were Western blotted with a murine anti-Myc mAb (1:5,000; Invitrogen). Blots were visualized with the WestPico Chemiluminescent kit (Pierce Chemical Co.).

Online supplemental material

The videos show the first two rounds of cell division in *C. elegans* embryos exposed to the RNAi treatments described in Fig. 7. Video 1 shows normal cell division in buffer-injected embryos. Video 2 shows normal cell division in *cpg-2(RNAi)* embryos. Video 3 shows that *cpg-1/cpg-2(RNAi)* embryos fail to complete the first cell division. Video 4 shows that *sqv-5(RNAi)* produces the same phenotypic defects as *cpg-1/cpg-2(RNAi)*. Online supplemental material is available at http://www.jcb.org/cgi/content/full/jcb.200603003/DC1.

We would like to thank Ian McLeod (The Scripps Research Institute) for helpful discussion regarding mass spectrometry; Lindsay Lewellyn, Amy Maddox, and Nathan Portier for assistance with worm imaging studies; and Raffi Aroian for many helpful and inspiring discussions.

This work was supported by National Institutes of Health grants GM33063 (to J. Esko) and RR11823 (to J. Yates) and Genetics Training Grant GM08666 (to S. Olson). K. Oegema is a Pew Scholar in the Biomedical Sciences and is supported by funding from the Ludwig Institute for Cancer Research.

Submitted: 1 March 2006
Accepted: 11 May 2006

References

Audhya, A., F. Hyndman, I.X. McLeod, A.S. Maddox, J.R. Yates III, A. Desai, and K. Oegema. 2005. A complex containing the Sm protein CAR-1 and the RNA helicase CGH-1 is required for embryonic cytokinesis in *Caenorhabditis elegans*. *J. Cell Biol.* 171:267–279.

Bame, K.J., and J.D. Esko. 1989. Undersulfated heparan sulfate in a Chinese hamster ovary cell mutant defective in heparan sulfate N-sulfotransferase. *J. Biol. Chem.* 264:8059–8065.

Beintema, J.J. 1994. Structural features of plant chitinases and chitin-binding proteins. *FEBS Lett.* 350:159–163.

Bellin, R., I. Capila, J. Lincecum, P.W. Park, O. Reizes, and M.R. Bernfield. 2002. Unlocking the secrets of syndecans: transgenic organisms as a potential key. *Glycoconj. J.* 19:295–304.

Berninsone, P., H.Y. Hwang, I. Zemtseva, H.R. Horvitz, and C.B. Hirschberg. 2001. SQV-7, a protein involved in *Caenorhabditis elegans* epithelial invagination and early embryogenesis, transports UDP-glucuronic acid, UDP-N-acetylgalactosamine, and UDP-galactose. *Proc. Natl. Acad. Sci. USA.* 98:3738–3743.

Brenner, S. 1974. The genetics of *Caenorhabditis elegans*. *Genetics.* 77:71–94.

Bulik, D.A., and P.W. Robbins. 2002. The *Caenorhabditis elegans sqv* genes and functions of proteoglycans in development. *Biochim Biophys Acta.* 1573:247–257.

Bulik, D.A., G. Wei, H. Toyoda, A. Kinoshita-Toyoda, W.R. Waldrip, J.D. Esko, P.W. Robbins, and S.B. Selleck. 2000. *sqv-3 -7*, and *-8*, a set of genes affecting morphogenesis in *Caenorhabditis elegans*, encode enzymes required for glycosaminoglycan biosynthesis. *Proc. Natl. Acad. Sci. USA.* 97:10838–10843.

Esko, J.D. 1993. Special considerations for proteoglycans and glycosaminoglycans and their purification. *In* Current Protocols in Molecular Biology. F. Ausubel, R. Brent, B. Kingston, D. Moore, J. Seidman, J. Smith, K. Struhl, A. Varki, and J. Coligan, editors. Greene Publishing and Wiley-Interscience, New York. 17.2.1–17.2.9.

Esko, J.D., and A. Manzi. 1996. Measurement of uronic acids. *In* Current Protocols in Molecular Biology. F. Ausubel, R. Brent, B. Kingston, D. Moore, J. Seidman, J. Smith, K. Struhl, A. Varki, and J. Coligan, editors. Greene Publishing and Wiley-Interscience, New York. 17.9.8–17.9.11.

Esko, J.D., and L. Zhang. 1996. Influence of core protein sequence on glycosaminoglycan assembly. *Curr. Opin. Struct. Biol.* 6:663–670.

Esko, J.D., and S.B. Selleck. 2002. Order out of chaos: Assembly of ligand binding sites in heparan sulfate. *Annu. Rev. Biochem.* 71:435–471.

Ghiselli, G., and R.V. Iozzo. 2000. Overexpression of bamacan/SMC3 causes transformation. *J. Biol. Chem.* 275:20235–20238.

Goldberg, M., D. Septier, O. Rapoport, R.V. Iozzo, M.F. Young, and L.G. Ameye. 2005. Targeted disruption of two small leucine-rich proteoglycans, biglycan and decorin, excerpts divergent effects on enamel and dentin formation. *Calcif. Tissue Int.* 77:297–310.

Grobe, K., and J.D. Esko. 2002. Regulated translation of heparan sulfate N-acetylglucosamine N-deacetylase/N-sulfotransferase isozymes by structured 5′-untranslated regions and internal ribosome entry sites. *J. Biol. Chem.* 277:30699–30706.

Habuchi, O. 2000. Diversity and functions of glycosaminoglycan sulfotransferases. *Biochim. Biophys. Acta.* 1474:115–127.

Hassell, J., Y. Yamada, and E. Arikawa-Hirasawa. 2002. Role of perlecan in skeletal development and diseases. *Glycoconj. J.* 19:263–267.

Herman, T., E. Hartwieg, and H.R. Horvitz. 1999. *sqv* mutants of *Caenorhabditis elegans* are defective in vulval epithelial invagination. *Proc. Natl. Acad. Sci. USA*. 96:968–973.

Hwang, H.Y., and H.R. Horvitz. 2002a. The *Caenorhabditis elegans* vulval morphogenesis gene *sqv-4* encodes a UDP-glucose dehydrogenase that is temporally and spatially regulated. *Proc. Natl. Acad. Sci. USA*. 99:14224–14229.

Hwang, H.Y., and H.R. Horvitz. 2002b. The SQV-1 UDP-glucuronic acid decarboxylase and the SQV-7 transporter may act in the Golgi apparatus to affect *Caenorhabditis elegans* vulval morphogenesis and embryonic development. *Proc. Natl. Acad. Sci. USA*. 99:14218–14223.

Hwang, H.Y., S.K. Olson, J.R. Brown, J.D. Esko, and H.R. Horvitz. 2003a. The *Caenorhabditis elegans* genes *sqv-2* and *sqv-6*, which are required for vulval morphogenesis, encode glycosaminoglycan galactosyltransferase II and xylosyltransferase. *J. Biol. Chem.* 278:11735–11738.

Hwang, H.Y., S.K. Olson, J.D. Esko, and H.R. Horvitz. 2003b. *Caenorhabditis elegans* early embryogenesis and vulval morphogenesis require chondroitin biosynthesis. *Nature.* 423:439–443.

Iozzo, R.V. 1998. Matrix proteoglycans: from molecular design to cellular function. *Annu. Rev. Biochem.* 67:609–652.

Izumikawa, T., H. Kitagawa, S. Mizuguchi, K.H. Nomura, K. Nomura, J. Tamura, K. Gengyo-Ando, S. Mitani, and K. Sugahara. 2004. Nematode chondroitin polymerizing factor showing cell-/organ-specific expression is indispensable for chondroitin synthesis and embryonic cell division. *J. Biol. Chem.* 279:53755–53761.

Jones, A.R., R. Francis, and T. Schedl. 1996. GLD-1, a cytoplasmic protein essential for oocyte differentiation, shows stage- and sex-specific expression during *Caenorhabditis elegans* germline development. *Dev. Biol.* 180:165–183.

Kamath, R.S., M. Martinez-Campos, P. Zipperlen, A.G. Fraser, and J. Ahringer. 2001. Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*. *Genome Biol.* 2:RESEARCH0002.1–RESEARCH0002.10.

Lee, M.H., and T. Schedl. 2001. Identification of in vivo mRNA targets of GLD-1, a maxi-KH motif containing protein required for *C. elegans* germ cell development. *Genes Dev.* 15:2408–2420.

Lee, M.H., and T. Schedl. 2004. Translation repression by GLD-1 protects its mRNA targets from nonsense-mediated mRNA decay in *C. elegans*. *Genes Dev.* 18:1047–1059.

MacCoss, M.J., W.H. McDonald, A. Saraf, R. Sadygov, J.M. Clark, J.J. Tasto, K.L. Gould, D. Wolters, M. Washburn, A. Weiss, et al. 2002. Shotgun identification of protein modifications from protein complexes and lens tissue. *Proc. Natl. Acad. Sci. USA*. 99:7900–7905.

Maddox, A.S., B. Habermann, A. Desai, and K. Oegema. 2005. Distinct roles for two *C. elegans* anillins in the gonad and early embryo. *Development.* 132:2837–2848.

Marneros, A.G., and B.R. Olsen. 2005. Physiological role of collagen XVIII and endostatin. *FASEB J.* 19:716–728.

Merzendorfer, H., and L. Zimoch. 2003. Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases. *J. Exp. Biol.* 206:4393–4412.

Mizuguchi, S., T. Uyama, H. Kitagawa, K.H. Nomura, K. Dejima, K. Gengyo-Ando, S. Mitani, K. Sugahara, and K. Nomura. 2003. Chondroitin proteoglycans are involved in cell division of *Caenorhabditis elegans*. *Nature.* 423:443–448.

Niisato, K., A. Fujikawa, S. Komai, T. Shintani, E. Watanabe, G. Sakaguchi, G. Katsuura, T. Manabe, and M. Noda. 2005. Age-dependent enhancement of hippocampal long-term potentiation and impairment of spatial learning through the Rho-associated kinase pathway in protein tyrosine phosphatase receptor type Z-deficient mice. *J. Neurosci.* 25:1081–1088.

Olson, S.K., and J.D. Esko. 2004. Proteoglycans. *In* Encyclopedia of Biological Chemistry. Vol. 3. W. Lennarz and M.D. Lane, editors. Elsevier, Oxford. 549–555.

Ozerdem, U., and W.B. Stallcup. 2004. Pathological angiogenesis is reduced by targeting pericytes via the NG2 proteoglycan. *Angiogenesis.* 7:269–276.

Reinke, V., H.E. Smith, J. Nance, J. Wang, C. Van Doren, R. Begley, S.J. Jones, E.B. Davis, S. Scherer, S. Ward, and S.K. Kim. 2000. A global profile of germline gene expression in *C. elegans*. *Mol. Cell.* 6:605–616.

Schwartz, N.B., and M. Domowicz. 2002. Chondrodysplasias due to proteoglycan defects. *Glycobiology.* 12:57R–68R.

Schwartz, N.B., and M. Domowicz. 2004. Proteoglycans in brain development. *Glycoconj. J.* 21:329–341.

Sugahara, K., T. Mikami, T. Uyama, S. Mizuguchi, K. Nomura, and H. Kitagawa. 2003. Recent advances in the structural biology of chondroitin sulfate and dermatan sulfate. *Curr. Opin. Struct. Biol.* 13:612–620.

Toyoda, H., A. Kinoshita-Toyoda, and S.B. Selleck. 2000. Structural analysis of glycosaminoglycans in *Drosophila* and *Caenorhabditis elegans* and demonstration that *tout-velu*, a *Drosophila* gene related to EXT tumor suppressors, affects heparan sulfate *in vivo*. *J. Biol. Chem.* 275:2269–2275.

Venegas, A., J.C. Goldstein, K. Beauregard, A. Oles, N. Abdulhayoglu, and J.A. Fuhrman. 1996. Expression of recombinant microfilarial chitinase and analysis of domain function. *Mol. Biochem. Parasitol.* 78:149–159.

Washburn, M.P., D. Wolters, and J.R. Yates III. 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* 19:242–247.

Washburn, M.P., R. Ulaszek, C. Deciu, D.M. Schieltz, and J.R. Yates III. 2002. Analysis of quantitative proteomic data generated via multidimensional protein identification technology. *Anal. Chem.* 74:1650–1657.

Wells, L., K. Vosseller, R.N. Cole, J.M. Cronshaw, M.J. Matunis, and G.W. Hart. 2002. Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications. *Mol. Cell. Proteomics.* 1:791–804.

Wolters, D.A., M.P. Washburn, and J.R. Yates III. 2001. An automated multidimensional protein identification technology for shotgun proteomics. *Anal. Chem.* 73:5683–5690.

Wright, H.T., G. Sandrasegaram, and C.S. Wright. 1991. Evolution of a family of N-acetylglucosamine binding proteins containing the disulfide-rich domain of wheat germ agglutinin. *J. Mol. Evol.* 33:283–294.

Wu, R.R., and J.R. Couchman. 1997. cDNA cloning of the basement membrane chondroitin sulfate proteoglycan core protein, bamacan: a five domain structure including coiled-coil motifs. *J. Cell Biol.* 136:433–444.

Yamada, S., I. Van Die, D.H. Van den Eijnden, A. Yokota, H. Kitagawa, and K. Sugahara. 1999. Demonstration of glycosaminoglycans in *Caenorhabditis elegans*. *FEBS Lett.* 459:327–331.

Yamada, S., Y. Okada, M. Ueno, S.S. Deepa, S. Nishimura, M. Fujita, I. Van Die, Y. Hirabayashi, and K. Sugahara. 2002. Determination of the glycosaminoglycan-protein linkage region oligosaccharide structures of proteoglycans from *Drosophila melanogaster* and *Caenorhabditis elegans*. *J. Biol. Chem.* 277:31877–31886.

Zelensky, A.N., and J.E. Gready. 2003. Comparative analysis of structural properties of the C-type-lectin-like domain (CTLD). *Proteins.* 52:466–477.